

OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/782,390

DATE: 04/26/2001
 TIME: 18:03:10

Input Set : N:\Crf3\RULE60\09782390.txt
 Output Set: N:\CRF3\04262001\I782390.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: LaBrie, Samuel T.
 8 Lal, Preeti
 9 Murry, Lynn E.

C--> 11 (ii) TITLE OF INVENTION: NOVEL TUBBY HOMOLOGUE

13 (iii) NUMBER OF SEQUENCES: 4

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 17 (B) STREET: 3174 Porter Dr.
 18 (C) CITY: Palo Alto
 19 (D) STATE: CA
 20 (E) COUNTRY: USA
 21 (F) ZIP: 94304

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Diskette
 25 (B) COMPUTER: IBM Compatible
 26 (C) OPERATING SYSTEM: DOS
 27 (D) SOFTWARE: FastSEQ for Windows Version 2.0

29 (vi) CURRENT APPLICATION DATA:

30 (A) APPLICATION NUMBER: US/09/782,390

31 (B) FILING DATE: 12-Feb-2001

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 08/812,824

35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J
 39 (B) REGISTRATION NUMBER: 36,749
 40 (C) REFERENCE/DOCKET NUMBER: PF-0232US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 415-855-0555
 44 (B) TELEFAX: 415-845-4166

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 491 amino acids
 50 (B) TYPE: amino acid
 51 (C) STRANDEDNESS: single
 52 (D) TOPOLOGY: linear

54 (ii) MOLECULE TYPE: peptide

56 (vii) IMMEDIATE SOURCE:

57 (A) LIBRARY: HNT2NOT01
 58 (B) CLONE: Consensus

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

62 Met Glu Ala Ser Arg Cys Arg Leu Ser Pro Ser Gly Asp Ser Val Phe
 63 1 5 10 15
 64 His Glu Glu Met Met Lys Met Arg Gln Ala Lys Leu Asp Tyr Gln Arg
 65 20 25 30

ENTERED

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/782,390

TIME: 18:03:10

Input Set : N:\Crif3\RULE60\09782390.txt

Output Set: N:\CRF3\04262001\I782390.raw

```

66 Leu Leu Leu Glu Lys Arg Gln Arg Lys Lys Arg Leu Glu Pro Phe Met
67      35      40      45
68 Val Gln Pro Asn Pro Glu Ala Arg Leu Arg Arg Ala Lys Pro Arg Ala
69      50      55      60
70 Ser Asp Glu Gln Thr Pro Leu Val Asn Cys His Thr Pro His Ser Asn
71      65      70      75      80
72 Val Ile Leu His Gly Ile Asp Gly Pro Ala Ala Val Leu Lys Pro Asp
73      85      90      95
74 Glu Val His Ala Pro Ser Val Ser Ser Ser Val Val Glu Glu Asp Ala
75      100      105      110
76 Glu Asn Thr Val Asp Thr Ala Ser Lys Pro Gly Leu Gln Glu Arg Leu
77      115      120      125
78 Gln Lys His Asp Ile Ser Glu Ser Val Asn Phe Asp Glu Glu Thr Asp
79      130      135      140
80 Gly Ile Ser Gln Ser Ala Cys Leu Glu Arg Pro Asn Ser Ala Ser Ser
81      145      150      155      160
82 Gln Asn Ser Thr Asp Thr Gly Thr Ser Gly Ser Ala Thr Ala Ala Gln
83      165      170      175
84 Pro Ala Asp Asn Leu Leu Gly Asp Ile Asp Asp Leu Glu Asp Phe Val
85      180      185      190
86 Tyr Ser Pro Ala Pro Gln Gly Val Thr Val Arg Cys Arg Ile Ile Arg
87      195      200      205
88 Asp Lys Arg Gly Met Asp Arg Gly Leu Phe Pro Thr Tyr Tyr Met Tyr
89      210      215      220
90 Leu Glu Lys Glu Glu Asn Gln Lys Ile Phe Leu Leu Ala Ala Arg Lys
91      225      230      235      240
92 Arg Lys Lys Ser Lys Thr Ala Asn Tyr Leu Ile Ser Ile Asp Pro Val
93      245      250      255
94 Asp Leu Ser Arg Glu Gly Glu Ser Tyr Val Gly Lys Leu Arg Ser Asn
95      260      265      270
96 Leu Met Gly Thr Lys Phe Thr Val Tyr Asp Arg Gly Ile Cys Pro Met
97      275      280      285
98 Lys Gly Arg Gly Leu Val Gly Ala Ala His Thr Arg Gln Glu Leu Ala
99      290      295      300
100 Ala Ile Ser Tyr Glu Thr Asn Val Leu Gly Phe Lys Gly Pro Arg Lys
101      305      310      315      320
102 Met Ser Val Ile Ile Pro Gly Met Thr Leu Asn His Lys Gln Ile Pro
103      325      330      335
104 Tyr Gln Pro Gln Asn Asn His Asp Ser Leu Leu Ser Arg Trp Gln Asn
105      340      345      350
106 Arg Thr Met Glu Asn Leu Val Glu Leu His Asn Lys Ala Pro Val Trp
107      355      360      365
108 Asn Ser Asp Thr Gln Ser Tyr Val Leu Asn Phe Arg Gly Arg Val Thr
109      370      375      380
110 Gln Ala Ser Val Lys Asn Phe Gln Ile Val His Lys Asn Asp Pro Asp
111      385      390      395      400
112 Tyr Ile Val Met Gln Phe Gly Arg Val Ala Asp Asp Val Phe Thr Leu
113      405      410      415
114 Asp Tyr Asn Tyr Pro Leu Cys Ala Val Gln Ala Phe Gly Ile Gly Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/782,390

DATE: 04/26/2001

TIME: 18:03:10

Input Set : N:\Crif3\RULE60\09782390.txt

Output Set: N:\CRF3\04262001\I782390.raw

```

115          420          425          430
116 Ser Ser Phe Asp Lys Arg Ile Gln Thr Leu Arg Met Gln Glu Leu Cys
117          435          440          445
118 Glu Leu His Arg Gln His His Ser Ala Ala Ser Leu Val His Arg Thr
119          450          455          460
120 Ala Cys Gln Arg Trp Val Gly His Pro Trp Arg Gln Leu Pro Gln Ser
121          465          470          475          480
W--> 122 Ser Leu Val Gly Pro Asp Leu Xaa Leu His Met
123          485          490
125 (2) INFORMATION FOR SEQ ID NO: 2:
127 (i) SEQUENCE CHARACTERISTICS:
128 (A) LENGTH: 1525 base pairs
129 (B) TYPE: nucleic acid
130 (C) STRANDEDNESS: single
131 (D) TOPOLOGY: linear
133 (ii) MOLECULE TYPE: cDNA
135 (vii) IMMEDIATE SOURCE:
136 (A) LIBRARY: HNT2NOT01
137 (B) CLONE: Consensus
139 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
141 GCACGAGGTG GGGGCTTTCC TCGGTGGCGG GCATGGAGGC TTCGCGCTGC CGGCTCAGTC 60
142 CCAGCGGCGA CAGTGTCTTC CATGAAGAAA TGATGAAGAT GCGACAGGCT AAGCTGGATT 120
143 ATCAGAGGCT ACTACTTGAG AAGAGGCAAA GGAAAAAGCG CCTTGAGCCA TTTATGGTGC 180
144 AGCCCAATCC AGAAGCCAGG CTACGTCGGG CAAAGCCAAG GGCCAGTGAT GAGCAGACTC 240
145 CCTTGGTGAA CTGTCTACTC CCCCACAGCA ATGTCATCTT ACATGGTATT GATGGTCCAG 300
146 CTGCTGTCCCT GAAACCCAGC GAAGTTCATG CTCCATCAGT AAGCTCCTCT GTTGTGGAAG 360
147 AAGATGCTGA AAACACCGTG GATACTGCTT CCAAGCCAGG ACTTCAGGAG CGTCTCCAAA 420
148 AGCATGATAT CTCTGAAAGT GTGAAGTTCG ATGAGGAGAC TGATGGAATA TCCCAGTCAG 480
149 CATGTTTAGA AAGACCCAAAT TCTGCATCAA GCCAGAATTC AACCGATACA GGCACCTCCG 540
150 GTTCTGCTAC TGCCGCCCAA CCAGCTGATA ACCTCCTGGG AGACATAGAC GACCTGGAGG 600
151 ACTTTGTGTA TAGTCCTGCC CCTCAAGGTG TCACAGTAAG ATGTCGGATA ATCCGGGATA 660
152 AAAGGGGAAT GGATCGGGGT CTCTTCCCCA CCTACTATAT GTACTTGGAA AAAGAAGAAA 720
153 ATCAGAAGAT ATTTCTTCTT GCAGCTAGAA AGCGGAAAAA GAGCAAAACA GCCAACTACC 780
154 TTATCTCCAT TGATCCAGTT GATTTATCTC GTGAAGGAGA AAGTTATGTC GGCAAGCTTA 840
155 GATCCAACCT CATGGGGACC AAGTTTACAG TTTATGACCG TGGCATCTGC CCCATGAAGG 900
156 GCCGGGGTTT GGTAGGAGCG GCCCACACCC GGCAGGAGCT GGCTGCCATC TCCTATGAAA 960
157 CAAACGTACT TGGATTAAAA GGTCTTAGGA AAATGTCTGT GATCATTCTT GGAATGACAC 1020
158 TGAATCATAA GCAGATCCCC TATCAGCCAC AAAACAACCA TGACAGTTTG CTCTCAAGGT 1080
159 GGCAGAACAG AACTATGGAA AATCTGGTTG AGCTGCACAA CAAGGCCCCC GTCTGGAACA 1140
160 GTGACACTCA GTCCTATGTC CTCAACTTCC GTGGCCGGGT CACTCAGGCG TCTGTGAAGA 1200
161 ACTTCCAGAT AGTCCACAAA AATGACCCTG ATTATATAGT CATGCAGTTT GGACGTGTGG 1260
162 CAGATGACGT GTTCACACTG GATTACAACT ACCCACTTTG TGCAGTACAG GCCTTTGGCA 1320
163 TCGGTCTTTC TAGCTTTGAC AAACGTATCC AAACCTTGAG AATGCAGGAG CTCTGTGAGC 1380
164 TCCACCGTCA GCACCATTCA GCTGCATCCC TTGTGCACAG GACTGCCTGC CAGCGTTGGG 1440
165 TGGGACACCC GTGGCGGCAG CTCCTCAGT CTTCCTTGT CGGCCCTGAC CTNTNACTAC 1500
166 ATATGTAGNA GCCCGAGACC AAAAA 1525
168 (2) INFORMATION FOR SEQ ID NO: 3:
170 (i) SEQUENCE CHARACTERISTICS:
171 (A) LENGTH: 505 amino acids

```

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/782,390

TIME: 18:03:10

Input Set : N:\Crif3\RULE60\09782390.txt

Output Set: N:\CRF3\04262001\I782390.raw

```

172      (B) TYPE: amino acid
173      (C) STRANDEDNESS: single
174      (D) TOPOLOGY: linear
176      (ii) MOLECULE TYPE: peptide
178      (vii) IMMEDIATE SOURCE:
179          (A) LIBRARY: GenBank
180          (B) CLONE: 1279766
182      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
184 Met Thr Ser Lys Pro His Ser Asp Trp Ile Pro Tyr Ser Val Leu Asp
185   1           5           10           15
186 Asp Glu Gly Ser Asn Leu Arg Gln Gln Lys Leu Asp Arg Gln Arg Ala
187           20           25           30
188 Leu Leu Glu Gln Lys Gln Lys Lys Arg Gln Glu Pro Leu Met Val
189           35           40           45
190 Gln Ala Asn Ala Asp Gly Arg Pro Arg Ser Arg Arg Ala Arg Gln Ser
191           50           55           60
192 Glu Glu Gln Ala Pro Leu Val Glu Ser Tyr Leu Ser Ser Ser Gly Ser
193           65           70           75           80
194 Thr Ser Tyr Gln Val Gln Glu Ala Asp Ser Ile Ala Ser Val Gln Leu
195           85           90           95
196 Gly Ala Thr Arg Pro Pro Ala Pro Ala Ser Ala Lys Lys Ser Lys Gly
197           100          105          110
198 Ala Ala Ala Ser Gly Gly Gln Gly Gly Ala Pro Arg Lys Glu Lys Lys
199           115          120          125
200 Gly Lys His Lys Gly Thr Ser Gly Pro Ala Thr Leu Ala Glu Asp Lys
201           130          135          140
202 Ser Glu Ala Gln Gly Pro Val Gln Ile Leu Thr Val Gly Gln Ser Asp
203          145          150          155          160
204 His Asp Lys Asp Ala Gly Glu Thr Ala Ala Gly Gly Gly Ala Gln Pro
205           165          170          175
206 Ser Gly Gln Asp Leu Arg Ala Thr Met Gln Arg Lys Gly Ile Ser Ser
207           180          185          190
208 Ser Met Ser Phe Asp Glu Asp Glu Asp Glu Asp Glu Asn Ser Ser Ser
209           195          200          205
210 Ser Ser Gln Leu Asn Ser Asn Thr Arg Pro Ser Ser Ala Thr Ser Arg
211           210          215          220
212 Lys Ser Ile Arg Glu Ala Ala Ser Ala Pro Ser Pro Ala Ala Pro Glu
213          225          230          235          240
214 Pro Pro Val Asp Ile Glu Val Gln Asp Leu Glu Glu Phe Ala Leu Arg
215           245          250          255
216 Pro Ala Pro Gln Gly Ile Thr Ile Lys Cys Arg Ile Thr Arg Asp Lys
217           260          265          270
218 Lys Gly Met Asp Arg Gly Met Tyr Pro Thr Tyr Phe Leu His Leu Asp
219           275          280          285
220 Arg Glu Asp Gly Lys Lys Val Phe Leu Leu Ala Gly Arg Lys Arg Lys
221           290          295          300
222 Lys Ser Lys Thr Ser Asn Tyr Leu Ile Ser Val Asp Pro Thr Asp Leu
223          305          310          315          320
224 Ser Arg Gly Gly Asp Ser Tyr Ile Gly Lys Leu Arg Ser Asn Leu Met

```

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/782,390

TIME: 18:03:10

Input Set : N:\Crif3\RULE60\09782390.txt

Output Set: N:\CRF3\04262001\I782390.raw

```

225          325          330          335
226 Gly Thr Lys Phe Thr Val Tyr Asp Asn Gly Val Asn Pro Gln Lys Ala
227          340          345          350
228 Ser Ser Ser Thr Leu Glu Ser Gly Thr Leu Arg Gln Glu Leu Ala Ala
229          355          360          365
230 Val Cys Tyr Glu Thr Asn Val Leu Gly Phe Lys Gly Pro Arg Lys Met
231          370          375          380
232 Ser Val Ile Val Pro Gly Met Asn Met Val His Glu Arg Val Cys Ile
233          385          390          395          400
234 Arg Pro Arg Asn Glu His Glu Thr Leu Leu Ala Arg Trp Gln Asn Lys
235          405          410          415
236 Asn Thr Glu Ser Ile Ile Glu Leu Gln Asn Lys Thr Pro Val Trp Asn
237          420          425          430
238 Asp Asp Thr Gln Ser Tyr Val Leu Asn Phe His Gly Arg Val Thr Gln
239          435          440          445
240 Ala Ser Val Lys Asn Phe Gln Ile Ile His Gly Asn Asp Pro Asp Tyr
241          450          455          460
242 Ile Val Met Gln Phe Gly Arg Val Ala Glu Asp Val Phe Thr Met Asp
243          465          470          475          480
244 Tyr Asn Tyr Pro Leu Cys Ala Leu Gln Ala Phe Ala Ile Ala Leu Ser
245          485          490          495
246 Ser Phe Asp Ser Lys Leu Ala Cys Glu
247          500          505
249 (2) INFORMATION FOR SEQ ID NO: 4:
251   (i) SEQUENCE CHARACTERISTICS:
252       (A) LENGTH: 506 amino acids
253       (B) TYPE: amino acid
254       (C) STRANDEDNESS: single
255       (D) TOPOLOGY: linear
257   (ii) MOLECULE TYPE: protein
259   (vii) IMMEDIATE SOURCE:
260       (A) LIBRARY: GenBank
261       (B) CLONE: 1305497
263   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
265 Met Thr Ser Lys Pro His Ser Asp Trp Ile Pro Tyr Ser Val Leu Asp
266   1      5      10      15
267 Asp Glu Gly Arg Asn Leu Arg Gln Gln Lys Leu Asp Arg Gln Arg Ala
268          20      25      30
269 Leu Leu Glu Gln Lys Gln Lys Lys Arg Gln Glu Pro Leu Met Val
270          35      40      45
271 Gln Ala Asn Ala Asp Gly Arg Pro Arg Ser Arg Arg Ala Arg Gln Ser
272          50      55      60
273 Glu Glu Gln Ala Pro Leu Val Glu Ser Tyr Leu Ser Ser Ser Gly Ser
274          65      70      75      80
275 Thr Ser Tyr Gln Val Gln Glu Ala Asp Ser Leu Ala Ser Val Gln Leu
276          85      90      95
277 Gly Ala Thr Arg Pro Thr Ala Pro Ala Ser Ala Lys Arg Thr Lys Ala
278          100     105     110
279 Ala Ala Thr Ala Gly Gly Gln Gly Gly Ala Ala Arg Lys Glu Lys Lys

```

VERIFICATION SUMMARY

DATE: 04/26/2001

PATENT APPLICATION: US/09/782,390

TIME: 18:03:11

Input Set : N:\Crf3\RULE60\09782390.txt

Output Set: N:\CRF3\04262001\I782390.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:11 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:122 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1